Journal of Biomedical Informatics 52 (2014) 212–221

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Use patterns of health information exchange through a multidimensional lens: Conceptual framework and empirical validation

Liran Politi [a,*], Shlomi Codish [b], Iftach Sagy [b], Lior Fink [a]

[a] Department of Industrial Engineering & Management, Ben-Gurion University of the Negev, Israel
[b] Clinical Research Center, Soroka University Medical Center, Israel

ARTICLE INFO

Article history:
Received 25 April 2014
Accepted 5 July 2014
Available online 14 July 2014

Keywords:
Use pattern
Health information exchange
Log file
Multidimensional analysis

ABSTRACT

Insights about patterns of system use are often gained through the analysis of system log files, which record the actual behavior of users. In a clinical context, however, few attempts have been made to typify system use through log file analysis. The present study offers a framework for identifying, describing, and discerning among patterns of use of a clinical information retrieval system. We use the session attributes of volume, diversity, granularity, duration, and content to define a multidimensional space in which each specific session can be positioned. We also describe an analytical method for identifying the common archetypes of system use in this multidimensional space. We demonstrate the value of the proposed framework with a log file of the use of a health information exchange (HIE) system by physicians in an emergency department (ED) of a large Israeli hospital. The analysis reveals five distinct patterns of system use, which have yet to be described in the relevant literature. The results of this study have the potential to inform the design of HIE systems for efficient and effective use, thus increasing their contribution to the clinical decision-making process.

1. Introduction

Health information exchange (HIE) systems are health information systems (HISs) that enable the electronic exchange and integration of patient-level health information across and within organizational boundaries [1,2]. HIE systems allow clinicians, regardless of their location and employers, to electronically exchange information about common patients [3]. This enables the data to "follow" the patient, allowing "re-use" of clinical data [1]. Potential advantages of these systems include improved quality of care and patient safety, cost reduction, and increased efficiency, e.g., [4,5].

Several studies have established that the extent to which these benefits are fulfilled depends on the implementation of the HIE system and its integration into clinicians' workflow [6]. Hence, shedding light on HIE usage patterns may promote the successful realization of their potential benefits. Such an analysis can address the components of the information technology (IT) productivity paradox – mismeasurement, mismanagement, and poor usability [7] – in the context of HISs.

Previous studies that empirically analyze patterns of HIE use are scarce. Prior studies measured organizational use as well as use made by individuals, which are the ultimate HIE end users and are therefore of great importance [8]. As the measurement of use at the individual level is difficult, several studies have turned to the analysis of system log files [2,9]. Nevertheless, using log files to explore healthcare processes is challenging due to their inherently complex, variable, and caregiver–contingent nature [10].

Individual usage patterns were usually characterized through the sequence, number, and types of screens viewed by HIE system users, while disregarding temporal traits, e.g., [10,11]. Another shortcoming of previous research is that the small number of studies that classified usage patterns generally used categories such as no use, basic use, or advanced use e.g., [2,12].

Against this backdrop, the present study offers a framework for identifying, describing, and discerning among patterns of use of a clinical information retrieval system. Although we focus our discussion on HIE systems, the proposed framework should be applicable for studying the use of electronic medical records (EMRs). In this framework, we suggest attributes that describe both the general and context-related use of the system and we account for temporal aspects of system use. We recommend a multilayered method of analysis to examine the attributes of system use and the associations among them.

The setting in which we empirically validate the proposed framework is the emergency medicine departments (EDs), specifically in the care for critically-ill patients. Effectively and correctly

* Corresponding author. Address: 1 Ben-Gurion Ave., Beer-Sheva 84105, Israel. Fax: +972 8 6472958.
E-mail address: liranpo@post.bgu.ac.il (L. Politi).

http://dx.doi.org/10.1016/j.jbi.2014.07.003
1532-0464/

diagnosing these patients in a timely manner is an important challenge physicians constantly face. These patients are typically not able to provide a medical history, and information is obtained from secondary sources such as family members and information systems. In this scenario, an HIE system is likely to have a significant impact on clinical decision making if information is readily accessible; the need for rapid decisions might render the scrutiny of an HIE system impractical.

We therefore chose to validate the framework with track log data from a widely-used HIE system in Israel, through which the medical data for over 50% of the Israeli population are accessible. To control for organizational and contextual factors, we focus on the use made by physicians within the busiest ED in Israel. Using the activity documentation of actual HIE users enables the identification of several multidimensional patterns of use. The main contribution of this study is in constructing and validating a framework for understanding HIE use. Gaining deep understanding of use patterns will assist in designing HIE systems to match clinical needs and incorporating the system into the decision-making process, thus enhancing its value to the clinicians and patients.

This article is organized as follows: in Section 2 we review past research on HIE patterns of use, explaining the motivation behind this stream of research. In Section 3 we outline a framework for describing patterns of use, drawing on the main attributes examined in previous studies within and outside the domain of health information. In Section 4 we apply the framework to HIE system logs that document use by physicians in an Israeli ED, demonstrating the meaningful conclusions that can be derived from using the proposed framework. In Section 5 we discuss this study's contributions and limitations, and offer avenues for future research.

2. Patterns of use of HIE systems

2.1. Research motivation

The potential advantages of HIE systems are not attained by merely implementing the systems. A two-level effort should guide the system implementation. First, a considerable effort should be made toward integrating the system into the users' workflow [6,12,13]. Nevertheless, studies have shown that placing the HIE system in the users' workflow does not necessarily lead users to fully utilize the system's abilities [2,9,14]. Significant effort should therefore also be placed on efficient use, which is contingent on system design. Successful implementation of an HIE system ought to include sufficient and directed endeavors that attend to these deficiencies [15]. Moreover, an examination of actual use may contribute to improving the system [9] and to estimating its impact on performance [16].

An exploration of HIE usage therefore involves the investigation of the system's integration into the workflow, followed by in-depth examinations of actual system use. Actual information system use can be defined as the act of accessing the system and utilizing its features in the end-user's workflow and can be stratified into several levels: individual, team, organizational, and inter-organizational [8]. We next briefly review the literature on these topics.

2.2. Use of HIE systems at the organizational level

In spite of the potential benefits of HIE systems, their adoption rates remain relatively low, although they are increasing. Whereas primary care practices present high adoption levels [17], inpatient care services are making their way more slowly [18]. Hospital implementation rates rose significantly in the past few years [19]. Hospitals with highly active EDs or those that maintain inter-organizational relationships are more likely to deploy HIE methods and systems [19,20].

HIE system users are diverse and pose different information needs [9,13]. Different users in the same organizational unit may integrate the system into their work patterns in different ways [21]. The organizational level of analysis may therefore miss out on important aspects of HIE use at the individual level.

The majority of studies on HIE systems use have focused on the individual level, specifically on the patterns of use exercised by physicians and nurses [8]. Despite the variance in the manner of use among individuals, some common role-based workflow patterns have been described by characteristics such as the timing of HIE use, types of accessed information, and main consumers of the retrieved data [21]. The notion of common workflow integration patterns supports the idea of common patterns of actual use.

2.3. Use patterns of HIE systems at the individual level

The individual level of analysis highlights the diversity of methods and measures for HIE system use. While some studies analyzed use patterns by observing and interviewing users [21], others utilized electronic log files of HIE systems, occasionally combined with semi-structured interviews [2,22–25]. Transaction log files contain documentation of electronic interactions between users and information retrieval systems [26], showing what information was displayed to the users at their request.

Studies that used log file analysis described patterns of use by means of measuring and analyzing the following indicators:

- *Types of users* that accessed the system, including physicians, nurses, administrative employees, and pharmacists [9,22], sometimes taking into account the additional variable of workplace (e.g., pediatric care, ambulatory care, hospital ED) [9,27].
- *Rate of patient encounters* in which the system was accessed [2,12,14,22].
- *Timing in relation to the encounter* in which the system was accessed (e.g., before the encounter, during the encounter, and retrospective use) [9,12].

Going beyond the dichotomous use/no-use approach applied in the previous measures is preferable when thoroughly describing the use of a system [9,28]. The following indicators delve into the attributes of each system access:

- *Time spent* per system access or per screen [23]. This temporal measure has been absent in most HIS use studies.
- *Diversity* and *types of accessed information* [9,24,27,29].
- *Frequency of basic access and of use of more advanced features* [2,14].
- *Sequence of screens* accessed by the user [9,11,23,25].

Although use patterns vary across users, the indicators summarized above have been employed to classify the use of HIE systems into relatively broad patterns. Two prominent classifications, which generally address the "breadth" of use, are:

- *No use, basic use, and advanced use*: These patterns were distinguished by the number and type of screens that were accessed during use. Basic access included a summary of patient history, lab results, and medications. The novel use pattern incorporated basic access with any additional views or inquiries [2,12,14].
- *Minimal (basic) use, repetitive searching, clinical information, mixed information, and demographic information*: This classification differs from the previous one mainly in the resolution by which novel use patterns were specified [9].

Several important findings emerge from the studies cited above. The access rate to HIE systems is generally low regardless of the type of user. Specifically, physicians have been reported to access available HIE systems in less than 10% of their encounters with patients, e.g., [13,22]. The nature of use is often basic and includes viewing the following items: a summary of patient demographic information, concise history of prior visits and hospitalizations, lab test results, and discharge summaries [9,14,21,22]. Commonly accessed non-summary data include clinician notes and various laboratory and radiology tests [21,22,24,27]. A previous study conducted in EDs found that a repetitive use pattern (i.e., cycling between patient search screen and table of most recent encounters) was less common, whereas the clinical use pattern, which includes in-depth queries of the patient's clinical information, was more frequent [9].

3. A framework for characterizing use patterns

There is a lack of research dedicated to understanding patterns of use of HISs in general, and of HIE systems in particular. This is especially the case with log-based studies. Many of the studies in this area have been based on a descriptive analysis [8] and usage patterns were usually characterized by the sequence, number, and types of screens viewed by users. Only a few studies described patterns of user behavior. Temporal measures (e.g., the duration of information display) were almost entirely unattended in previous analyses.

We propose a framework that captures the richness of individual use, accounting for the various dimensions that might be relevant to understanding the diversity of actual use. This framework is designed to guide the analysis of electronic log files that document the use of health information. Since research on HIS use is not as developed as in other domains, the framework is based on a broader scope of studies that explored information system and website use. We later apply the framework to data on HIE use.

Log file analysis gained momentum in various areas, including information searching on the internet, information retrieval [30], and consumerism and e-commerce [31]. Information service operators have realized that large quantities of data are of no less importance than easy, quick, and efficient access [32]. Hence, large click-stream data repositories were analyzed to extract user access patterns [31,33] with the purpose of web personalization, business intelligence discovery, and user characterization and classification [33].

This method enables inductive inference based on empirical observations [30]. Advocates of this method assert that it facilitates the high-resolution study of individual-level activities and preferences on large-scale fields [34]. The method also allows decomposing decision making or information inquiry processes into steps, rather than considering final results alone.

3.1. Level of analysis

Log files can be segmented at different levels, resulting in different kinds of insights. They consist of *click-streams*, defined as "the sequence of pages, i.e., screens, that are accessed by a user" [31], p. 95. A *session* is defined as the click-stream created by a single user's activity within the system, during the time period between login and logout [25,31].

In addition to the session level, other prevalent levels of analysis include (1) single screen level, (2) data categories, and (3) sets of consecutive screens [31,33]. These three levels practically decompose sessions into smaller components, and they can be analyzed with simple statistics such as frequencies, means, and correlations. Complete session analysis, however, provides a more comprehensive view of user-system interaction, and thus it often expands to classifying user behaviors into usage cohorts [31,33]. We consequently chose the session level of analysis for our framework.

3.2. Session attributes

A session is often described by a set of quantifiable attributes based on its composition. A sequence of viewed screens is hence converted into a single multidimensional set of attributes, serving as an empirical observation. This representation aims at a rich, yet parsimonious, account of sessions that can be more easily analyzed statistically, enabling descriptive and predictive analyses [34].

Sessions are sometimes represented as a binary vector of all screens present in a session or all possible transitions between screens [35,36]. These two options result in a "flat" description that vastly increases the vector's dimension, causing overrepresentation of certain types of information that appear in a larger assortment of screens. This structure also poses a challenge for various clustering techniques [37]. Therefore, many studies turn to a more holistic and aggregative approach to session representation.

In the following section, we identify several attributes that were frequently employed to describe sessions in previous studies.

3.2.1. Volume

A common characteristic of a session is the amount of information or extent of user-system interaction it encompasses. This attribute defines the volume of user-system interaction and, as such, it may indicate the user's interest in the provided information [38]. Measurement of this attribute is often accomplished by counting the number of screens in a session [9,38–40].

3.2.2. Diversity

This attribute refers to the *different* types of information units displayed within a session, indicating the variety of the explored information [38]. Sessions may include large amounts of homogeneous information, and thus defined as high volume and low diversity. The session's level of diversity can be measured by the number or percentage of *different* information units that were displayed, e.g., different screens or domains viewed in a session [9,34,38,41,42].

3.2.3. Granularity

The previous attributes are based on the implicit premise that all units of information (i.e., screens) are essentially equivalent. Such a premise, however, is inaccurate because different units often represent information at different levels of granularity. While one screen may contain summarized information, for instance a list of laboratory tests and their results, another screen may display specific information, for example one blood test in detail. Correspondingly, a session may contain access to summarized information or to information of high granularity. An additional session attribute is thus the level of specificity of the information included in it.

Measuring this attribute often requires the classification of screens into a compatible hierarchy [34,43]. A session may be characterized according to the maximal level of specificity presented in any of its screens. Another possibility is to count the screens that pertain to each level of specificity [34]. This concept was also applied by Eason et al. [41], who examined the depth of access to e-journals, ranging from groups of journals to a full-text article in a specific journal.

3.2.4. Duration of screen display (DSD)

The dimension of time advances the ability to understand session behavior. Addressing the time dedicated to the entire session or to a specific screen introduces a dynamic dimension to the way sessions are described. Furthermore, information viewing time may attest to the level of importance the user attributes to the viewed information. Time-related traits of a session have been demonstrated to be linked to the user's level of expertise regarding the explored information [38] and to the type of information-seeking task [34,44]. This attribute may therefore indirectly reflect user perceptions and consequently provide a partial remedy to a prevalent criticism, that log-file analysis relates only to users' actions [26].

Established measurements of this aspect are the total amount of time a user dedicates to the entire session or the average time a screen is viewed [34,38,39]. As the distribution of the average may be skewed, it is sometimes replaced by the median [45].

3.2.5. Content

The previous attributes disregard the subject of the information units within the session. Using the contents of viewed screens may enhance the ability to assess the user's intentions or goals for accessing the system [34]. Moreover, this attribute complements information variety measures by referring to the actual content of the viewed screens and not just to their quantity. Using the session's content as an attribute has been shown to contribute to typifying HIE system use [9] and to predicting user decisions [34,40,46].

In order to more concisely describe session content, screens are sometimes categorized into types [9,34,39,40]. The quantification of session content may include binary indicators for whether information categories were accessed [43], or counting the number of screens that belong to each category [34,40].

The attributes of volume, diversity, granularity, duration, and content represent a multidimensional space in which each specific session can be defined. To facilitate the understanding of user behavior, it is necessary to group the multitude of different sessions into a narrow set of clusters that represent archetypical patterns of user behavior, homogenously distributed within clusters and heterogeneously distributed across clusters. Such an approach is frequently adopted in organizational research for the purpose of clustering organizations based on their attributes [47,48].

Fig. 1 graphically summarizes the suggested framework. The following section represents an empirical validation of the proposed framework with HIE log files. While not being a formal part of the framework, the methods and procedures employed as part of the empirical analysis may also serve as guidelines for other applications of the framework.

4. Empirical validation

4.1. Use of HIE systems in the emergency department (ED)

EDs are dynamic, high-paced environments with increasing annual visit rates and potentially life-threatening situations that require quick responses [49]. Patient information deficiencies at the place and time of care are common [13,50] and have been identified as the cause of many healthcare quality and safety issues in EDs [24]. These gaps commonly include deficiencies in medical history and laboratory test results [50].

A longitudinal inspection of the patient's health condition and history upon arrival to an ED is imperative to an efficient care plan [24]. Although HIE use is clearly not the solution to all sources of medical uncertainty, it may support physicians in dealing with unfamiliar or unknown patients and with task complexity caused by the patient's characteristics or history [14]. These circumstances, along with the expected benefits of HIE, account for the belief that care in EDs may benefit the most from efficient HIE [2,13,24].

Most emergency physicians recognize the contribution of HIE systems to their departments, hospitals, and their patients' safety [13,15,51]. Nonetheless, available HIE systems are traditionally utilized by a minority of clinicians and for few encounters with patients [14]. Inhibitors of HIE use are often the acuity of the patient's condition, incomplete HIE data, and usability issues [8,13]. This gap between the potential benefits of HIE in EDs and its extent of use underscores the need to explore the patterns of its actual use.

4.2. Research setting

We apply the proposed framework to analyze patterns of use of the OFEK HIE system (dbMotion, Israel) in a large Israeli hospital. The log file being analyzed documents all information retrieval actions that were conducted in the system by physicians during the treatment of 1001 critically-ill patients in a busy internal ED in a three-year period (2010–2012). Following the process depicted in Fig. 1, we present the context of system use, the stages of segmenting the raw log file into sessions that are characterized by multiple attributes, and the application of descriptive and clustering statistical techniques to attain meaningful insights about system use patterns.

4.2.1. OFEK HIE system

Soroka University Medical Center (SUMC) serves over 1.1 million people as the only tertiary medical center in the southern area of Israel (60% of the land area of the country). SUMC's ED is the busiest in Israel with over 210,000 visits annually. SUMC is owned by Clalit Health Services, which operates 14 hospitals and over 1300 primary care clinics in Israel and is the primary health care provider for over half the population of the country.

In 2005, a web-interface-based federated model HIE system named OFEK was implemented at SUMC. OFEK retrieves data from several dispersed Clalit HISs databases and from most large hospitals in Israel and transmits these data to the point of care (e.g., SUMC ED). The system presents the information to the user as an integrated patient file. The obtained information includes medical history, previous hospital admissions, community and outpatient clinics, discharge letters, laboratory and radiograph results, and medications and prescriptions.

A few studies have examined the attitude toward and skills of using OFEK [52], the actual use of OFEK, and its contribution to the healthcare process [27,29,53]. The study by Ben-Assuli et al. [29], for example, significantly linked viewing certain types of information in a session with admission decisions. These studies did not aim at typifying use patterns and therefore did not explore complete sessions.

4.2.2. Emergency services in SUMC

Critically-ill patients presenting to SUMC ED are treated in SUMC's resuscitation room (RR), which is equipped with advanced medical equipment and operates only when a critical treatment is required. Our empirical validation focuses on internal medicine patients who were treated in the RR.

Some of the critical patients first arrive to the internal medicine ED, where they are registered and triaged by a nurse, and their condition deteriorates while in the ED. Others arrive to the RR directly by emergency medical services. Critical patients are often uncommunicative, thus rendering personal questioning impossible. At the time of the validation no EMR was used at SUMC ED, and relevant clinical information was therefore available only through the HIE system, the patient's paper chart, or family members. Patient charts contain the following data: treatment given by paramedics,

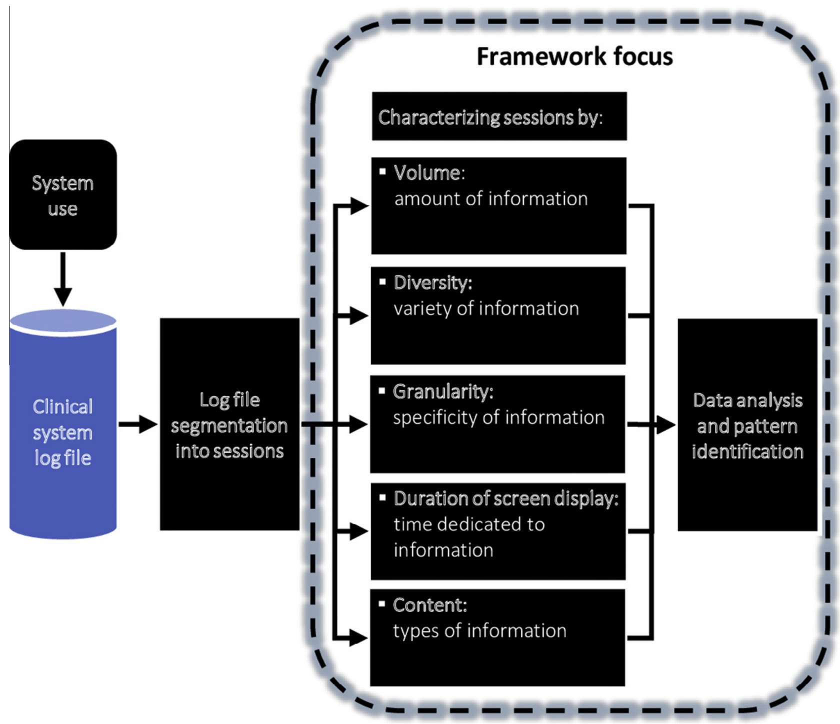

Fig. 1. An illustration of the proposed framework.

vital signs measured in the ED and RR, treatment given in the RR (including time, dosage, and way of medication), and emergency procedures the patient underwent. Historical data, however, can be obtained only by using the HIE system. After being treated in the RR, surviving patients are admitted either to an intensive care unit or to a relevant ward.

4.2.3. OFEK use characteristics

The HIE system can only be accessed by a username and a password. While logged-in, users are able to access patient data independent of the time of the encounter with the patient, depending on the clinician's role and authorizations. The system runs a 30-min idle automatic logout mechanism, intended to enforce authorizations. Access to data on a specific patient via the HIE system can be gained in two ways. The first is searching the patient's "medical record" using the patient's identifying details. Gaining access this way directs the user to an integrative "Patient data summary" screen, which displays the 3–4 most recent data items of each of the following: hospital admissions, prescribed medications, diagnoses given in hospitals and in community clinics, and lab results (an indicator for results received in the preceding week). The second way of access is locating the patient in the patient roster. This screen displays all patients who are currently registered in the ED or RR, their time of admission, and times of last results of laboratory tests and imaging scans performed after admission to the ED. Once the sought entry is found, the user may proceed to "Patient data summary" or directly to the latest laboratory test result or imaging scan.

A fixed side menu in the HIE system individual patient display enables users to navigate and change screens. Each screen pertains to one of five information categories: general, test results, procedures, documents, and previous visits. Once access to a specific patient file is achieved, any screen can be accessed either directly or via one intermediary screen, making the system almost entirely "flat". Patient-specific screens can be categorized into two levels, according to the specificity of data they display and navigational properties. *Level A screens* are those that display summarized data. *Level B screens* are those that display data of high specificity, such as specific lab test results or an account of a specific visit. Table 1 presents examples for screens of each category and level.

4.3. Dataset

The sample for this study included all adult patients who were treated in the SUMC RR between January 1, 2010 and December 31, 2012. A total of 1001 patients had 1051 RR admissions during this period. After obtaining IRB approval, a list of these patients was obtained. Our focus on cases in which patients were in need of urgent internal medicine services was expected to control for the confounding effects of situational variables.

The analysis is based on log files that were extracted from the HIE system databases. These files contain documentation of all information retrieval activities that were conducted by physicians only and for the specific patients in our sample. Each row in the log file represents access to a single HIE system screen, described by the following five attributes: patient ID, unique encounter ID, user ID, name of accessed screen, and time of access (Table 2a). The log file does not include within-screen activity (e.g., scrolling), documentation of unsuccessful data retrieval efforts, and timestamps for end of sessions.

We distinguish between two possible operational definitions of sessions. A *patient session* (P session) is the click-stream generated by the use of the system for a specific patient during the patient's specific visit to the ED. A single P session may consist of click-streams produced by several users. A *patient–user session* (P–U session) is the click-stream generated by the activity of a specific user, for a specific patient, on a specific encounter [9]. The literature includes studies that focus on P sessions [2,10,12,14], on the user point of view [11,25], and on P–U sessions [9]. Table 2a presents an example of a single P session, which consists of two P–U sessions.

Table 1
Information available via OFEK by categories and levels.

Information category	Level A	Level B
General	Patient data summary, demographic details	Prescribed medications, medication allergies
Test results	Test results summary, imaging tests summary	Specific blood test result (e.g., biochemistry), specific imaging test (e.g. chest X-ray)
Procedures	Past procedures summary (surgeries, cardiac catheterizations)	Procedure notes
Documents	Discharge letters summary	Specific discharge letter
Previous visits	Outpatient clinic visits summary, previous ED visits	Description of a specific visit to an outpatient clinic, specific hospitalization details

Table 2
Example of log-file data conversion.

(2a) Format of raw data (before conversion)

Encounter ID	Patient ID	User ID	Screen name	Time
4	123	456	Patient data summary	4/9/2012 11:01:00
4	123	456	Document list	4/9/2012 11:01:05
4	123	456	Patient data summary	4/9/2012 11:10:08
4	123	456	Laboratory summary	4/9/2012 11:10:15
4	123	456	Biochemistry result	4/9/2012 11:10:22
4	123	789	Patient data summary	4/9/2012 12:15:23
4	123	789	Laboratory summary	4/9/2012 12:15:26

(2b) Format of final data (after conversion)

P–U session ID	Volume	Diversity	Median DSD	Granularity	Content – information categories				
					General	Test results	Procedures	Documents	Previous visits
1	5	4	7	1	1	1	0	1	0
2	2	2	3	0	1	1	0	0	0

The final data set contained 16,715 screen views performed by 238 physicians. The system was used during 84.97% of the admissions (893 RR admission). These 893 P sessions consisted of 1661 P–U sessions, implying an average of 1.86 P–U sessions per P session. Because over 50% of P sessions included use of the system by two or more physicians, we opted for P–U sessions as the unit of analysis. Consistent with the literature on information system use, this choice is based on the premise that use patterns are contingent on the individual characteristics of the user, implying that screen views by different physicians cannot be treated as representing a single physician.

Consequently, the raw dataset of screen views was converted into records of P–U sessions, which consisted of the following attributes:

- *Volume:* the number of screen views in a session [9,38,39]. Two views of the same screen in a single session are considered two different views.
- *Diversity*: the number of *unique* screens that were accessed [9,41]. This attribute considers two views of the same screen as a single unique screen.
- *Granularity*: the level of the most specific screen viewed in a session. This attribute is a binary variable, coded as 1 if the session included a Level B screen and as 0 if the session included only Level A screens.
- *Content – information categories*: five binary variables that indicate (coded as 1) if access was made to five data categories - general, test results, procedures, documents, and previous visits.
- *Median duration of screen display (Median DSD)*: the median display time of a screen in a session. We calculated the display time for a screen as the time difference, in seconds, between time of access to the screen and time of access to its consecutive screen. We chose to represent the dimension of time for each session by computing the median DSD. The median was selected as a way of mitigating the effect of long, yet unreliable, durations of screen display [45]. This choice reduced the skewness of DSD distribution from 5.45 for average values to 3.90 for median values.

To determine session endings, previous research has used a threshold that determines the maximal screen display time for which it is reasonable to believe the user has not left the system [31,33]. We applied this concept of a threshold to mitigate the effect of DSD outliers. Several physicians we interviewed stated that, based on their experience, a physician views an HIE system screen for no longer than 30 s. Therefore, a more rigorous threshold of 60 s was implemented, meaning that any screen display that lasted for longer than 60 s was truncated to 60 s. While this procedure significantly influenced average median DSD values (mean decreased from 13.46 to 7.65 and standard deviation decreased from 102.07 to 7.94), it had little influence on median DSD values themselves, affecting only those that were higher than 60 s (2.3% of the P–U sessions).

Table 2a presents two P–U sessions in their raw form and Table 2b demonstrates the final processed form of these sessions.

4.4. Data analysis and results

4.4.1. Descriptive analysis

We used SPSS 20 for data analysis. The descriptive statistics for the nine P–U session attributes are presented in Table 3. The table shows that an average session is composed of a sequence of 10.06 screens (volume), representing 4.85 different screens (diversity). On average, therefore, screens are viewed 2.07 times in each session. The mean median DSD is 7.65 s and its distribution is right-skewed, as is the distribution of session volume. As the other six attributes are binary, the means are in fact the proportion of sessions that were assigned with a value of 1. Most sessions (68%) incorporate at least one granular (Level B) screen. While the

Table 3
Descriptive statistics of P–U sessions.

	Volume	Diversity	Median DSD (sec.)	Granularity	Content – information categories				
					General	Test results	Procedures	Documents	Previous visits
Mean	10.06	4.85	7.65	0.68	0.88	0.72	0.01	0.61	0.09
Median	7	5	5.5	1	1	1	0	1	0
Std. Dev.	10.15	2.13	7.94	0.47	0.33	0.45	0.11	0.49	0.29
Minimum	1	1	0	0	0	0	0	0	0
Maximum	83	15	60	1	1	1	1	1	1
Skewness	3.83	0.66	3.90	–	–	–	–	–	–

general (88%), test results (72%), and documents (61%) information categories are included in most sessions, the procedures (1%) and previous visits (9%) information categories are seldom included.

Pearson and Spearman bivariate correlation coefficients are displayed in Table 4. Interestingly, viewing test results is negatively correlated with viewing general information ($\rho = -0.24$, $p < 0.01$) and documents ($\rho = -0.25$, $p < 0.01$).

4.4.2. Cluster analysis

As the goal of our empirical validation was to identify common patterns of use based on session attributes, a grouping technique was required. Consistent with the literature, we applied cluster analysis techniques to cluster sessions into "usage clusters" [31,33]. Cluster analysis is a dominant technique for item grouping, as it enables classification based on rich descriptions through multiple and diverse variables [54]. It is frequently applied when data reduction and taxonomy description are required [32,33] and is considered useful for extracting behavior patterns from "noisy" data [10]. This analysis assumes a holistic approach and characterizes usage by considering all attributes simultaneously.

Because all attributes had either symmetric binary or ratio scales, they were analyzed jointly [55]. For the purpose of clustering, attribute values were standardized to Z scores to avoid over-representing attributes with wider ranges [54]. The cluster analysis was conducted in two stages, as recommended by Hair et al. [54]. First, a hierarchical cluster analysis was performed to determine the number of clusters and their centroids. Sessions were clustered by a hierarchical average linkage method. The measure of dissimilarity (i.e., the distance between cluster centroids) was the squared Euclidean distance, prevalent in centroid-based analyses. An agglomeration coefficient schedule was used to select the cluster solution, indicating that sessions would be best classified into five clusters. Second, in order to validate and refine the hierarchical five-cluster solution, the centroids of these clusters were used as input (seed points) to a non-hierarchical K-means cluster analysis.

The final cluster solution is presented in Table 5 and Fig. 2. While Table 5 presents the cluster centroids (i.e., the means of sessions for each cluster) in terms of raw attribute values, Fig. 2 presents the cluster centroids in terms of standardized attribute values, to allow their graphical comparison. The final cluster solution included two relatively large clusters, together encompassing over 87% of the sessions, whereas the remaining sessions were covered by three relatively small clusters. Analysis of variance (ANOVA) is commonly used to determine that a specific variable distinguishes among clusters [54]. Using this procedure showed that all nine attributes significantly distinguished among clusters at the 0.05 level.

The results show that *cluster A*, which includes 60.02% of the sessions, is characterized by relatively low mean values in all attributes, with volume and diversity of 5.23 and 3.82 screens, respectively. The median DSD of 5.08 s is also the lowest among the five clusters. *Cluster E,* the smallest of all clusters (1.45%), also presents low mean values in most attributes, but its median DSD (52.88 s) is the highest of all clusters. *Cluster D*, in contrast, is the most voluminous (48.62 screens), diverse (8.28 screens), and granular (1.00) among all clusters. This cluster includes sessions with relatively high access rates to all information categories. *Cluster B*, the second largest cluster, encompasses sessions of high volume (17.67 screens) and granularity (0.99), with relatively high access rates to general and test results information. Finally, *cluster C* is similar to cluster A in terms of volume (6.47 screens) and diversity (4.51 screens), yet its median DSD (21.26 s) is about four times that of cluster A.

4.5. Use patterns

The rate of using the HIE system in our research setting, as reflected by the log data, is higher than previously published for ED patients. Upon treating a critically-ill patient, a physician tends to access 7–10 screens. The diversity of screens is often limited, implying a repetitive or concise access pattern [9]. System use is often quite rapid, with a few seconds dedicated to each screen. Dwelling on a particular screen for longer than 8 s is quite rare. The frequently basic and rapid use of the HIE system can be

Table 4
Correlation matrix.

	Volume	Diversity	Median DSD	Granularity	Content – information categories			
					General	Test results	Procedures	Documents
Diversity[P]	0.69[b]	–	–	–	–	–	–	–
Median DSD[P]	−0.03	−0.02	–	–	–	–	–	–
Granularity[S]	0.68[b]	0.66[b]	0.19[b]	–	–	–	–	–
General[S]	0.24[b]	0.44[b]	0.01	−0.06[a]	–	–	–	–
Test Results[S]	0.58[b]	0.54[b]	0.20[b]	0.63[b]	−0.24[b]	–	–	–
Procedures[S]	0.07[b]	0.11[b]	0.07[b]	0.02	0.04	−0.01	–	–
Documents[S]	0.20[b]	0.35[b]	0.16[b]	−0.13[b]	0.48[b]	−0.25[b]	0.06[a]	–
Previous Visits[S]	0.14[b]	0.25[b]	0.17[b]	0.08[b]	0.12[b]	−0.02	0.04	0.09[b]

[a] Correlation is significant at the 0.05 level (2-tailed).
[b] Correlation is significant at the 0.01 level (2-tailed).
[P] Coefficients in this row are Pearson coefficients.
[S] Coefficients in this row are Spearman coefficients.

Table 5
Cluster centroids – means of attribute values for five-cluster solution.

Cluster	A	B	C	D	E
Number of Sessions	995	457	135	47	24
Percentage	60.02%	27.56%	8.14%	2.83%	1.45%
Mean attribute value					
Volume	5.23	17.67	6.47	48.62	4.96
Diversity	3.82	6.89	4.51	8.28	3.21
Median DSD	5.08	6.78	21.26	6.29	52.88
Granularity	0.54	0.99	0.56	1.00	0.46
Content – information categoriesl					
General	0.84	0.96	0.88	1.00	0.54
Test results	0.58	0.99	0.67	1.00	0.79
Procedures	0.01	0.02	0.01	0.04	0.00
Documents	0.54	0.73	0.73	0.87	0.46
Previous visits	0.05	0.14	0.18	0.19	0.13

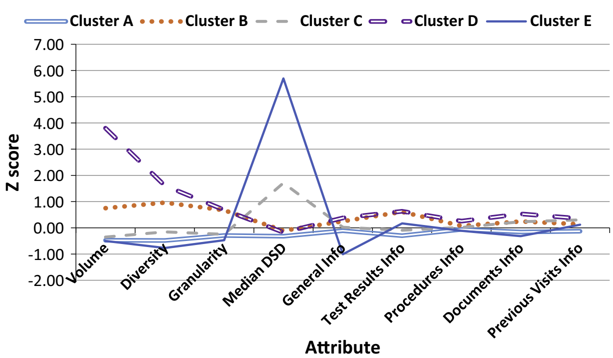

Fig. 2. Cluster profiles – means of standardized attribute values for five-cluster solution.

explained by the commonly urgent nature of RR treatments. Such an environment may induce basic HIE system use [2].

The vast majority of P–U sessions include access to general summarized information and test results. Past documents are also widely accessed. Viewing rates of previous visits and procedures are considerably lower than those of the other three information categories. This finding conforms to conclusions from previous studies done on OFEK HIE system [56]. One explanation is that some of the information in these categories (e.g., visits to outpatient clinics) may be less relevant to critical care. A second, more plausible explanation is that a great deal of information regarding past visits and procedures is provided via test results or general summary screens. These findings are consistent with previous studies, showing that laboratory test results, discharge summaries, and information summaries are of great importance to treatment, particularly in the ED [9,14,21,22,24,56].

An additional descriptive finding is that access time is not associated with the volume or diversity of consumed information, but rather with the type of information. Access to information of high granularity, as well as access to test results, procedures, documents, and previous visits is associated with reduced navigation speed. As these information types are often highly detailed, prolonged viewing is reasonable.

Exploring past documents is associated with viewing general information, indicating that screens on these information categories cater to similar or complementary information needs. Examining test results is negatively associated with viewing general information and documents, suggesting that these screens support different information needs.

Finally, the cluster analysis identified five archetypical patterns of use. We call the most prevalent pattern (cluster A) *"quick and basic"* because it involves fast access to a small number of screens. Such sessions often involve only summary screens that present general, test results, and documents information (46% of the sessions in this cluster). This pattern is similar in essence to the basic usage patterns identified in previous studies [2,9,14].

The infrequent *"slow and basic"* pattern (cluster E) differs from the former pattern primarily in the dimension of time, having a median DSD that is 10 times longer. Furthermore, this pattern involves viewing more information about test results and previous visits, but less general information and documents, suggesting a need to invest more cognitive effort in understanding the case at hand.

The *"broad and deep"* pattern (cluster D), which is also relatively infrequent, includes access to large volumes of information (48.62 screens on average). This pattern includes viewing highly-specific information and it stands out as a pattern that involves the highest viewing rate of information in all five categories. This type of use is perhaps reasonable for long treatments, complex medical conditions, and abundance of available information in the system.

The *"quick inquiry of many lab results"* is the second most prevalent pattern (cluster B). While the volume and diversity of consumed information in this pattern are intermediate, this pattern stands out in its high rate of access to general information and test results, which are accessed quickly.

The *"slow inquiry of previous visits and documents"* pattern (cluster C) is similar to the "quick and basic" and "slow and basic" patterns in terms of session volume and diversity. However, this pattern is characterized by longer durations and higher inspection rates of previous visits and documents.

The holistic perspective, emerging from our framework, produced insights that could not be obtained from bivariate analyses. Our findings demonstrate that it is the combination of multiple attributes that distinguishes among different patterns of system use, rather than a single attribute or linear relationships among attributes.

5. Discussion

5.1. Contributions and implications

This study extends recent findings on patterns of using HISs in general and HIE systems in particular based on the analysis of system log files that record the actual behavior of users. To the best of our knowledge, this study is the first to clearly delineate a framework for analyzing patterns of HIS use sessions. We do so by defining the major session attributes and suggesting a multi-attribute analysis of complete sessions. These attributes encompass general use characteristics such as the volume of information, its diversity, and speed of access, as well as content-related characteristics such as granularity and specific information categories. Our framework expands on previous research, which typically included no analysis of temporal elements of system use [14,22], by adding the dimension of time in identifying distinguishable patterns of use. We demonstrate the value of all these attributes in distinguishing among different use patterns. The patterns of use emerging from the data, following the implementation of the proposed framework, represent an expansion of existing classifications of system use [2,14].

The proposed framework is not limited to the specific attributes analyzed in this study. These attributes were chosen based on a review of the general literature on log-file analysis and were defined as broad concepts (e.g., volume, diversity, granularity, and content) that reflect the breadth and depth of information accessed by users. The specific operationalization of attributes depends on the objectives of the analysis and on the availability of data. Moreover, the framework suggests a specific multivariate method of analysis, independent of attribute definitions, that does not have to be applied if the objectives are different. Although our approach is not complex, it produces novel insights.

Previous studies on usage patterns analyzed a pooled log file, recording the behavior of various types of users [9,12]. The present study also innovates by analyzing and classifying sessions that are limited to physicians, controlling for the variance introduced by including additional medical staff.

Information displays on HISs are commonly uniform for all users and for all patients. The identification of different and distinguishable patterns of use highlights the potential value of system customization. Identifying information units that are frequently jointly accessed can help in determining what information should be displayed, when, and how [32]. It can also help in determining the preferred mode of presentation given typical durations of screen displays.

An important conclusion of our empirical analysis is that use of the system for critically-ill patients tends to be concise. This finding amplifies the effect of displaying information items simultaneously, particularly the default screens ("gateway screens") displayed to the user upon entering a patient record. A possible scheme is to create patient profile screens, in addition to existing screens, uniquely designed for specific clinical conditions. Another approach could be to create shortcut links that connect various screens in different information categories, further "flattening" the system. These links may also change dynamically as the session unfolds and more can be inferred about its characteristics. Such online, dynamic classification can rely on the findings provided by clustering techniques.

Such personalization should be carried out carefully to avoid the routinization of use patterns that are common but not necessarily optimal. This problem can be prevented by comparing actual patterns with treatment guidelines and, consequently, designing the system to facilitate the practice of treatment protocols. Reviewing the information inquiry process may also highlight information disparities and system deficiencies.

The analysis of usage patterns could help address the sources of the IT productivity paradox in health care – mismeasurement, mismanagement, and poor usability [7]. Log-based patterns of use provide management with an objective tool for measuring and supervising actual practice. Customizing the system to align with guidelines and user needs may enhance its usability and, consequently, improve the quality of care.

5.2. Limitations and future research

Information use research based on activity logs has inherent limitations. While logs contain documentation of accessed information, there is no indication of whether the information that was displayed was required or used in the decision making process [30]. Moreover, the log file analyzed in this study provided no indication on the availability of information. Users may not attempt to access certain screens because they are able to see, from a summary screen for example, that there is no relevant information available. Consequently, short use patterns may be caused by a lack of information and not by intentional basic use of the system.

The findings concerning the five specific use patterns have limited external validity, and caution should be exercised in generalizing them to other medical settings. The data analyzed in this study represent user activity for a specific HIE system, characterized by a specific architecture, user interface, and data sources, which may influence user behavior. Moreover, as no EMR was available at the ED in our research setting, the HIE system operated as the only electronic source of historical clinical data about patients. The system use was recorded in a specific medical setting that involves critically-ill patients. While this approach is likely to increase homogeneity, it poses a challenge to generalizing the findings to other settings and medical conditions.

To address these limitations, future research should revalidate the framework presented in this study with additional sets of data. Such studies are advised to use the log files of HIE systems because of their ability to provide a broader view of users' information needs based on the integration of data from multiple sources. To exploit the potential of the proposed framework, the log file being analyzed should include, at minimum, identifying information about users and patients, as well as the timestamps of screen views. If possible, the log file should indicate the data available at the time of access, shown to affect patterns of use [24], and within-screen activity (e.g., scrolling). Such additional data can facilitate the understanding of user behavior by highlighting the gaps between the data available in the system and the data of interest to the user.

Careful consideration should be practiced regarding the manner in which attributes are measured, taking into account such criteria as data availability, level of analysis, and desired pattern resolution. The "curse of dimensionality", which underscores the tradeoff between richness of description and difficulty of analysis [37], should also be taken into account.

The study of usage patterns is motivated by the notion that the full realization of the potential benefits of HIE depends on its effective integration into the care-giving process [6,12]. While this study aims at identifying patterns of use, it does not consider the context in terms of situational, patient, and user characteristics. This shortcoming calls for additional research to investigate the antecedents of usage patterns, such as the medical setting, patient clinical condition, and user experience.

Another key objective of HIE implementation efforts is improving performance in the care-giving process. A recommended path for additional research would be to ascertain whether the manner in which the system is used is associated with various performance measures for quality of care and resource utilization [16]. Better understanding of the antecedents and consequents of HIS use is expected to enhance the contribution of these systems to the quality of medical care.

References

[1] Hersh W. A stimulus to define informatics and health information technology. BMC Med Inform Decis Mak 2009;9(1):24. http://dx.doi.org/10.1186/1472-6947-9-24.
[2] Vest JR, Jasperson J, Zhao H, Gamm LD, Ohsfeldt R. Use of a health information exchange system in the emergency care of children. BMC Med Inform Decis Mak 2011;11(1):78. http://dx.doi.org/10.1186/1472-6947-11-78.
[3] Blumenthal D, Glaser JP. Information technology comes to medicine. N Engl J Med 2007;356(24):2527–34.
[4] Kaelber DC, Bates DW. Health information exchange and patient safety. J Biomed Inform 2007;40(6 Suppl):S40–5. http://dx.doi.org/10.1016/j.jbi.2007.08.011.
[5] Chaudhry B, Wang J, Wu S, Maglione M, Mojica W, Roth E, et al. Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. Ann Intern Med 2006;144(10):742–52.
[6] Frisse ME, Holmes RL. Estimated financial savings associated with health information exchange and ambulatory care referral. J Biomed Inform 2007;40(6 Suppl):S27–32. http://dx.doi.org/10.1016/j.jbi.2007.08.004.
[7] Jones SS, Heaton PS, Rudin RS, Schneider EC. Unraveling the IT productivity paradox – lessons for health care. N Engl J Med 2012;366(24):2243–5. http://dx.doi.org/10.1056/NEJMp1204980.
[8] Vest JR, Jasperson J. What should we measure? conceptualizing usage in health information exchange. J Am Med Inform Assoc 2010;17(3):302–7. http://dx.doi.org/10.1136/jamia.2009.000471.
[9] Vest JR, Jasperson J. How are health professionals using health information exchange systems? measuring usage for evaluation and system improvement. J Med Syst 2012;36(5):3195–204. http://dx.doi.org/10.1007/s10916-011-9810-2.
[10] Rebuge Á, Ferreira DR. Business process analysis in healthcare environments: a methodology based on process mining. Inform Syst 2012;37(2):99–116. http://dx.doi.org/10.1016/j.is.2011.01.003.

[11] Chen ES, Cimino JJ. Patterns of usage for a Web-based clinical information system. In: MEDINFO: Proceedings of the 11th world congress on medical informatics; 2004. p. 18–22.

[12] Vest JR, Gamm LD, Ohsfeldt RL, Zhao H, Jasperson J. Factors associated with health information exchange system usage in a safety-net ambulatory care clinic setting. J Med Syst 2011;36(4):2455–61. http://dx.doi.org/10.1007/s10916-011-9712-3.

[13] Shapiro JS, Kannry J, Kushniruk AW, Kuperman G. The New York clinical information exchange (NYCLIX) clinical advisory subcommittee, Kuperman GJ. emergency physicians' perceptions of health information exchange. J Am Med Inform Assoc 2007;14(6):700–5. http://dx.doi.org/10.1197/jamia.M2507.

[14] Vest JR, Zhao H, Jasperson J, Gamm LD, Ohsfeldt RL. Factors motivating and affecting health information exchange usage. J Am Med Inform Assoc 2011;18(2):143–9. http://dx.doi.org/10.1136/jamia.2010.004812.

[15] Rudin R, Volk L, Simon S, Bates D. What affects clinicians' usage of health information exchange? Appl Clin Inform 2011;2(3):250–62. http://dx.doi.org/10.4338/ACI-2011-03-RA-0021.

[16] Devaraj S, Kohli R. Performance impacts of information technology: is actual usage the missing link? Manag Sci 2003;49(3):273–89. http://dx.doi.org/10.1287/mnsc.49.3.273.12736.

[17] Fontaine P, Ross SE, Zink T, Schilling LM. Systematic review of health information exchange in primary care practices. J Am Board Fam Med 2010;23(5):655–70. http://dx.doi.org/10.3122/jabfm.2010.05.090192.

[18] Jha AK, Doolan D, Grandt D, Scott T, Bates DW. The use of health information technology in seven nations. Int J Med Inform 2008;77(12):848–54. http://dx.doi.org/10.1016/j.ijmedinf.2008.06.007.

[19] Furukawa MF, Patel V, Charles D, Swain M, Mostashari F. Hospital electronic health information exchange grew substantially In 2008–12. Health Aff 2013;32(8):1346–54. http://dx.doi.org/10.1377/hlthaff.2013.0010.

[20] Vest JR. More than just a question of technology: factors related to hospitals' adoption and implementation of health information exchange. Int J Med Informatics 2010;79(12):797–806. http://dx.doi.org/10.1016/j.ijmedinf.2010.09.003.

[21] Unertl KM, Johnson KB, Lorenzi NM. Health information exchange technology on the front lines of healthcare: workflow factors and patterns of use. J Am Med Inform Assoc 2012;19(3):392–400. http://dx.doi.org/10.1136/amiajnl-2011-000432.

[22] Johnson KB, Unertl KM, Chen Q, Lorenzi NM, Nian H, Bailey J, et al. Health information exchange usage in emergency departments and clinics: the who, what, and why. J Am Med Inform Assoc 2011;18(5):690–7. http://dx.doi.org/10.1136/amiajnl-2011-000308.

[23] Xiaowei L, Yuan X, Malin B. Towards understanding the usage pattern of web-based electronic medical record systems. In: World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2011 IEEE International Symposium; 2011, p. 1–7. doi: 10.1109/WoWMoM.2011.5986195.

[24] Hripcsak G, Sengupta S, Wilcox A, Green RA. Emergency department access to a longitudinal medical record. J Am Med Inform Assoc 2007;14(2):235–8. http://dx.doi.org/10.1197/jamia.M2206.

[25] Chen ES, Cimino JJ. Automated discovery of patient-specific clinician information needs using clinical information system log files. In: AMIA annual symposium proceedings; 2003: American Medical Informatics Association, p. 145–9.

[26] Peters TA. The history and development of transaction log analysis. Libr Hi Tech 1993;11(2):41–66. http://dx.doi.org/10.1108/eb047884.

[27] Nirel N, Rosen B, Sharon A, Blondheim O, Sherf M, Samuel H, et al. The impact of an integrated hospital-community medical information system on quality and service utilization in hospital departments. Int J Med Inform 2010;79(9):649–57. http://dx.doi.org/10.1016/j.ijmedinf.2010.06.005.

[28] Burton-Jones A, Straub DW. Reconceptualizing system usage: an approach and empirical test. Inform Syst Res 2006;17(3):228–46. http://dx.doi.org/10.1287/isre.1060.0096.

[29] Ben-Assuli O, Shabtai I, Leshno M. The influence of EHR components on admission decisions. Health Technol 2012;3(1):1–7. http://dx.doi.org/10.1007/s12553-013-0039-6.

[30] Jansen BJ. Search log analysis: what it is, what's been done, how to do it. Libr Inform Sci Res 2006;28(3):407–32. http://dx.doi.org/10.1016/j.lisr.2006.06.005.

[31] Cooley R. The use of web structure and content to identify subjectively interesting web usage patterns. ACM Trans Internet Technol 2003;3(2):93–116. http://dx.doi.org/10.1145/767193.767194.

[32] Pierrakos D, Paliouras G, Papatheodorou C, Spyropoulos CD. Web usage mining as a tool for personalization: a survey. User Model User Adap 2003;13(4):311–72. http://dx.doi.org/10.1023/A:1026238916441.

[33] Srivastava J, Cooley R, Deshpande M, Tan PN. Web usage mining: discovery and applications of usage patterns from web data. ACM SIGKDD Expl Newslett 2000;1(2):12–23. http://dx.doi.org/10.1145/846183.846188.

[34] Moe WW. Buying, searching, or browsing: differentiating between online shoppers using in-store navigational clickstream. J Consum Psychol 2003;13(1–2):29–39. http://dx.doi.org/10.1207/S15327663JCP13-1&2_03.

[35] Paliouras G, Papatheodorou C, Karkaletsis V, Spyropoulos CD. Clustering the users of large web sites into communities. In: Proceedings of the international conference on machine learning (ICML), Stanford, California, USA; 2000.

[36] Ishikawa H, Ohta M, Watanabe T, Yokoyama S, Katayama K. Toward active web usage mining for page recommendation and restructuring. In: Proceedings of I-KNOW 2003; Graz, Austria, p. 492–499.

[37] Donoho DL. High-dimensional data analysis: the curses and blessings of dimensionality. AMS Math Challenges Lect 2000:1–32.

[38] White RW, Dumais ST, Teevan J. Characterizing the influence of domain expertise on web search behavior. In: Proceedings of the second ACM international conference on web search and data mining. ACM; 2009. p. 132–41. http://dx.doi.org/10.1145/1498759.1498819.

[39] Benevenuto F, Rodrigues T, Cha M, Almeida V. Characterizing user behavior in online social networks. In: Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference. ACM; 2009. p. 49–62. http://dx.doi.org/10.1145/1644893.1644900.

[40] Van den Poel D, Buckinx W. Predicting online-purchasing behaviour. Eur J Oper Res 2005;166(2):557–75. http://dx.doi.org/10.1016/j.ejor.2004.04.022.

[41] Eason K, Richardson S, Yu L. Patterns of use of electronic journals. J Doc 2000;56(5):477–504. http://dx.doi.org/10.1108/EUM0000000007124.

[42] Huang CY, Shen YC, Chiang IP, Lin CS. Characterizing Web users' online information behavior. J Am Soc Inf Sci Technol 2007;58(13):1988–97. http://dx.doi.org/10.1002/asi.20669.

[43] Baglioni M, Ferrara U, Romei A, Ruggieri S, Turini F. Preprocessing and mining web log data for web personalization. AI*IA 2003: Advances in Artificial Intelligence. In: Proceedings of the 8th congress of the Italian association for artificial intelligence, Pisa, Italy, 2003. Berlin Heidelberg: Springer; 2003. p. 237–49.

[44] Kellar M, Watters C, Shepherd M. A field study characterizing Web-based information-seeking tasks. J Am Soc Inf Sci Technol 2007;58(7):999–1018. http://dx.doi.org/10.1002/asi.20590.

[45] Nicholas D, Huntington P, Williams P. Establishing metrics for the evaluation of touch screen kiosks. J Inform Sci 2001;27(2):61–71. http://dx.doi.org/10.1177/016555150102700201.

[46] Li S, Liechty JC, Montgomery AL. Modeling category viewership of web users with multivariate count models. Tepper School of Business 2002; Paper 334.

[47] Ketchen DJ, Combs JG, Russell CJ, Shook C, Dean MA, Runge J, et al. Organizational configurations and performance: a meta-analysis. Acad Manage J 1997;40(1):223–40. http://dx.doi.org/10.2307/257028.

[48] Fiss PC. A set-theoretic approach to organizational configurations. Acad Manage Rev 2007;32(4):1180–98. http://dx.doi.org/10.5465/AMR.2007.26586092.

[49] Institute of Medicine. Hospital-based emergency care: at the breaking point. Washington DC, USA: National Academy Press; 2006.

[50] Stiell A, Forster AJ, Stiell IG, van Walraven C. Prevalence of information gaps in the emergency department and the effect on patient outcomes. Can Med Assoc J 2003;169(10):1023–8.

[51] Ancker JS, Edwards AM, Miller MC, Kaushal R. Consumer perceptions of electronic health information exchange. Am J Prev Med 2012;43(1):76–80. http://dx.doi.org/10.1016/j.amepre.2012.02.027.

[52] Brainin E, Gilon G, Meidan N, Mushkat Y. The impact of intranet integrated patient medical file (IIPMF) assimilation on the quality of medical care and organizational advancements. The Israel National Institute for Health Policy and Health Services; 2005.

[53] Ben-Assuli O, Leshno M. Efficient use of medical IS: diagnosing chest pain. J Enterp Inform Manage 2012;25(4):413–23. http://dx.doi.org/10.1108/17410391211245865.

[54] Hair JF, Black WC, Babin BJ, Anderson RE, Tatham RL. Multivariate data analysis. 6th ed. New Jersey: Prentice Hall Upper Saddle River; 2005.

[55] Kaufman L, Rousseeuw PJ. Introduction. Finding groups in data: an introduction to cluster analysis. Wiley-Interscience; 2009. p. 3–49.

[56] Nirel N, Rosen B, Sharon A, Samuel H, Yair Y, Cohen AD, et al. Ofek virtual medical records: an evaluation of an integrated hospital-community online medical information system. Smokler Center for Health Policy Research; 2010.